

A MOLAR THEORY OF REINFORCEMENT SCHEDULES

HOWARD RACHLIN¹

STATE UNIVERSITY OF NEW YORK AT STONY BROOK

Behavior of subjects exposed to concurrent and individual interval and ratio schedules of reinforcement may be described in terms of a set of expressions relating the value of responses to their durations, a feedback equation relating reinforcement to response duration, and the assumption that subjects allocate their time among various responses so as to maximize value.

Key words: molar, reinforcement, economics, feedback, concurrent schedules, interval schedules, ratio schedules

Behavior of subjects exposed to various concurrent and individual schedules of reinforcement may be measured and described in several ways. This paper considers the implications of one system of measurement and description. It has the following characteristics:

1. Behavior is measured temporally. (Baum and Rachlin, 1969; Premack, 1965). That is, time (rather than rate) of response or reinforcement is the measure of dependent and independent variables. Time measures may be converted directly to rate measures if the duration of each response is known. The advantage of time measures is that the time of one response and another are commensurate and conversion constants, otherwise necessary, can be eliminated from equations relating one response to another.

2. Only three classes of behavior are observed and measured: I, the instrumental response; C, the contingent response (or reinforcement); N, all other behavior, considered as a single variable. The three classes are arranged to be mutually exclusive. In the usual operant situation with pigeons, for instance, the pigeon cannot peck a key (I) and consume grain (C) at the same time. The other behavior (N) is exclusive by definition.

3. A schedule of reinforcement is characterized by a feedback function: $C = f(I)$, which describes the reinforcement obtained as responding is varied (Baum, 1973).

4. All responses have values. The value of a response is a function of its duration and the value of a group of responses is a function

of the values of the responses comprising the group. Thus,

$$V_C = f(C)$$

$$V_I = f(I)$$

$$V_N = f(N)$$

and

$$V_{C, I, N} = f(V_C, V_I, V_N).$$

The Feedback Function

Although it might seem as though a feedback function (the function relating responding to the reinforcement obtained by that responding) should be a fixed characteristic of a schedule (an independent variable), in fact it is not completely independent of behavior. The reason for this is that the feedback function as defined above applies to molar, temporal aspects of behavior, whereas most schedules of reinforcement are arranged to be sensitive to a molecular aspect—*i.e.*, an individual response measured as an instantaneous event. Thus, the feedback function depends on the interresponse-time (IRT) distribution upon which the schedule acts. To complicate the analysis further, there is some evidence that the IRT distribution changes in the course of training. At the beginning of training, with a variable-interval schedule, the IRT distribu-

¹The author thanks John Gibbon and John Staddon for many helpful comments and criticisms. Preparation of the paper was supported by a grant from the National Science Foundation. Reprints may be obtained from Howard Rachlin, Psychology Department, State University of New York, Stony Brook, New York 11794.

tion and the distribution of response probabilities ("IRTs per opportunity"), usually reflect random responding. But at asymptote, these distributions clearly deviate from those that would be obtained if responding were random in time (Anger, 1956, McGill, 1963, Shimp, 1967).

Regardless of the "actual" asymptotic function, we can still examine the molar characteristics of various schedules and select a feedback function that would produce these characteristics. With respect to variable-interval schedules, the feedback function should have the following characteristics:

- (a) When responding is zero, reinforcement should be zero.
- (b) As responding increases, reinforcement should increase monotonically.
- (c) As responding approaches its maximum, reinforcement should approach the level programmed by the variable-interval schedule.
- (d) The rate of increase of reinforcement with responding should be a decreasing function of the level of responding. In other words, as responding increases from zero, reinforcement should increase rapidly to a high proportion of its maximum. But the rate of increase should thereafter decrease.

A simple function that has these properties is the power function:

$$C = aI^m, \quad (1)$$

where I is the fraction of the session spent responding and can vary from zero to 1.0. With variable-interval schedules, $0 < m < 1$. Other functions may also fit the requirements enumerated here. Baum (1973) proposed a hyperbolic-type function. Staddon and Motherall (*in press*) used a function based on random responding, to show that matching is consistent with maximizing. The present paper considers only the feedback function of Equation 1. Appendix I contains an attempt to derive the feedback function from an assumed burst-and-pause response distribution. The derived function is shown to be approximated closely by Equation 1. Also in Appendix I is an empirical feedback function consistent with Equation 1.

Figure 1 shows the feedback function of Equation 1 plotted with various values of m .

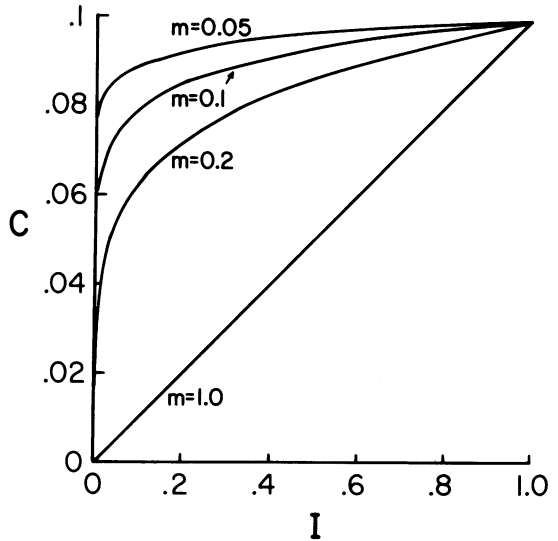


Fig. 1. The function, $C = aI^m$ where $a = 0.1$ and I varies from 0 to 1.0. The parameter is m .

All functions start at zero and increase to a value of $C = a$ when $I = 1.0$ (its maximum). Thus, a is the level of reinforcement arranged by the variable-interval schedule. When $0 < m < 1.0$, the function is negatively accelerated, providing the required decrease in slope with increasing I for variable-interval schedules.

The extreme points of m are meaningful. At $m = 0$, $C = a$ for all levels of I . Thus, at this extreme, the variable-interval schedule becomes a variable-time schedule with reinforcement delivered freely at level, a . At $m = 1$, $C = aI$ which is the feedback function for variable-ratio schedules. (We do not at this point attempt to account for behavior with fixed schedules of intervals, ratios, or times). The constant, a , is the ratio of reinforcement to responding imposed by the schedule. Thus, variation of a single parameter, m , of Equation 1 changes the schedule from variable time (at $m = 0$) to variable interval ($0 < m < 1.0$) to variable ratio (at $m = 1$). Variable-interval schedules with $m \approx 0$ would be more like variable-time schedules, while those with $m \approx 1$ would be more like variable-ratio schedules.

With Equation 1, one variable-interval schedule may differ from another only by the multiplicative constant, a , its programmed level of reinforcement. Thus, the form of the feedback function may remain constant regardless of the particular schedule. It is an open question whether the exponent m changes from one schedule to another. We as-

sume that m is usually constant across various variable-interval schedules but may be altered by altering the distribution of programmed reinforcements or by such devices as imposing a changeover delay (a delay between changing over from one response to the other and the soonest possible reinforcement) with concurrent variable-interval schedules. See Appendix I for a further discussion of these points.

A few examples will clarify the application of Equation 1 to actual schedules of reinforcement. With variable-ratio (VR) schedules, C , the contingent response, is directly proportional to I , the instrumental response. In Equation 1, such proportionality is obtained with $m = 1$. For instance, with a VR 10 schedule, 10 responses are required, on average, per reinforcement. Suppose each response takes 0.25 sec and reinforcements are 4 sec in duration. Then, with $m = 1$ in Equation 1 $a = C/I = 4/(0.25 \times 10) = 1.6$ and $C = 1.6I$.

With a variable-interval (VI) schedule, $0 < m < 1$. We assume here, in accordance with common experimental practice, that a session is composed only of responding (I) and not responding (N). The time spent consuming reinforcement (C) is considered to occur outside of the session. Thus, $I + N = \text{session duration}$. In a unit session, $I + N = 1$. Assume for a moment that with a variable-interval schedule a subject spends the entire session responding. In that case, $I = 1$ and, according to Equation 1, $C = a$, regardless of the value of m . With a variable-interval schedule, if a subject responds continuously during the session, it will receive all reinforcements exactly when they are programmed by the schedule. In most actual experiments with discrete responses, the programming equipment (the pulse former that converts responses into discrete pulses) sets a limit on response duration, and the subject can never obtain all programmed reinforcers, although obtained reinforcement approaches programmed reinforcement as rate of response increases to high values. But with continuous responding as studied by Baum and Rachlin (1969), for instance, reinforcers are simply delivered while a response is being made. If a subject spends the entire session responding it receives all programmed reinforcers. The constant, a , represents the programmed rate of reinforcement and equals the amount obtained (C) when the entire unit session is spent responding ($I = 1$).

When a subject spends less than the entire session responding ($0 < I < 1$), the *obtained* reinforcement (C) will be less than a . For instance, with each reinforcement available for 4 sec and a VI 40-sec schedule, $a = 4/40 = 0.1$ and $C = 0.1I^m$. In the illustrative cases of Figure 1 the constant, a , is set equal to 0.1.

Occasionally, the obtained reinforcement (C) may be greater than unity. As noted above, time consuming reinforcement is assumed to lie outside of the session. A continuous reinforcement schedule ($CRF = VR\ 1$) might provide 4 sec of eating for each response. Still assuming 0.25 sec per response, $a = C/I = 4/0.25 = 16$. If the subject spent half of the session responding ($I = 0.5$) then $C = 16(0.5) = 8$. The cases where $C > 1$ correspond to cases where the subject spends more time consuming reinforcement (outside the session) than the duration of the session ($I + N = 1$) itself.

Value

Rachlin and Burkhard (1978) have examined several value functions. One set of functions, called "power-function theory" was helpful in explaining some data with human subjects. We will consider this set of functions here:

$$\left. \begin{array}{l} V_C = C^x \\ V_I = I^y \\ V_N = N^z \end{array} \right\}, \quad (2)$$

and

$$V_{C, I, N} = V_C + V_I + V_N. \quad (3)$$

The exponents x , y , and z are independent of the exponent, m . That is, the power-function theory of value is independent of the assumption that the feedback function is a power function. Equation 2 could be true and Equation 1, false or *vice-versa*.

In addition to Equations 1, 2, and 3, we assume that asymptotic behavior during any schedule of reinforcement is allocated among the three responses, C , I , and N so as to maximize their collective value.

Application to Concurrent Schedules

Although concurrent schedules are, in principle, more complicated than the individual schedules of which they are composed, we consider them first because, with concurrent schedules, we may confine our attention to relative rather than absolute rates of response (Herrn-

stein, 1961). This allows several simplifications and illustrates well how the model operates.

With concurrent schedules, a session is divided into three parts: responding on schedule-1 (I_1), responding on schedule-2 (I_2), and not responding (N). The reinforcers, C_1 and C_2 , occur outside of the session. The value of these responses together, according to Equations 2 and 3 is:

$$V = (C_1 + C_2)^x + (I_1 + I_2)^y + (N)^z. \quad (4)$$

Substituting $C = aI^m$ (Equation 1) and $N = 1 - I_1 - I_2$ (a unit session);

$$V = (a_1 I_1^{m_1} + a_2 I_2^{m_2})^x + (I_1 + I_2)^y + (1 - I_1 - I_2)^z.$$

We now can take partial derivatives of the above expression with respect to I_1 and I_2 :

$$\begin{aligned} \frac{\partial V}{\partial I_1} &= x(a_1 I_1^{m_1} + a_2 I_2^{m_2})^{x-1} \cdot [a_1 m_1 I_1^{m_1-1}] \\ &\quad + y(I_1 + I_2)^{y-1} - z(1 - I_1 - I_2)^{z-1} \\ \frac{\partial V}{\partial I_2} &= x(a_1 I_1^{m_1} + a_2 I_2^{m_2})^{x-1} \cdot [a_2 m_2 I_2^{m_2-1}] \\ &\quad + y(I_1 + I_2)^{y-1} - z(1 - I_1 - I_2)^{z-1}. \end{aligned}$$

Any point of maximum value (we are assuming subjects allocate time to I_1 , I_2 , and N so as to maximize value) will be where:

$$\frac{\partial V}{\partial I_1} = \frac{\partial V}{\partial I_2} = 0.$$

The two partial derivatives are identical except for a single term in each (the term in brackets). These terms must equal each other at a maximum:

$$\begin{aligned} a_1 m_1 I_1^{m_1-1} &= a_2 m_2 I_2^{m_2-1} \\ \frac{I_1^{1-m_1}}{I_2^{1-m_2}} &= \frac{m_1}{m_2} \cdot \frac{a_1}{a_2}. \end{aligned} \quad (5)$$

Substituting $a = CI^{-m}$ (from Equation 1) in Equation 5:

$$\frac{I_1}{I_2} = \frac{m_1}{m_2} \cdot \frac{C_1}{C_2}. \quad (6)$$

Equation 6 is not necessarily a maximum solution of Equation 4. Rather, it describes a set of points, among which any maximum must lie. Thus, if behavior maximizes value it must conform to Equation 6. But if behavior conforms to Equation 6 it need not maximize value. (It can be shown, however, by techniques similar to the above, that Equation 6 is the only maximum if $I_1 + I_2$ is assumed constant and behavior maximizes $C_1 + C_2$ only.)

With concurrent variable-interval (VI VI) or variable-ratio (VR VR) schedules, where $m_1 \approx m_2$, Equation 6 is Herrnstein's matching

law (Herrnstein, 1970), which has been shown to hold for a wide variety of concurrent schedules (deVilliers, 1977). Thus, with concurrent schedules, when responding is distributed between alternatives so as to maximize overall reinforcement, matching will result. This relationship was shown stochastically (with computer simulation of various response distributions between alternatives) by Rachlin *et al.* (1976). The computer was programmed to distribute "responses" in a certain proportion (but otherwise randomly) to two variable-interval schedules. N was assumed to equal zero and the session divided between I_1 and I_2 . The proportion was varied from 0:100 through 50:50 to 100:0 per cent in steps of 10%. Reinforcements that would have been obtained with each proportion from each alternative were determined. It was found that the maximum total reinforcement (from both alternatives together) was obtained when the proportion of responses matched the proportion of reinforcements. Equation 6 confirms this relationship analytically.

In the computer simulation, changeover delays (delays imposed between switching from one alternative to another and the soonest possible reinforcement following the switch) were varied in duration. Higher changeover delays tended to flatten the feedback functions. Figure 1 shows that flatter feedback functions correspond to higher values of m . Appendix I shows that the slower alternation caused by the changeover delay results in flatter feedback functions. But the simulation by Rachlin *et al.* (1976) resulted in flatter feedback functions with a COD, even when response patterns remained constant. Thus, flatter feedback functions (higher values of m) may result both from the COD itself and from the alteration of response patterning that the COD causes. Let us say that a given changeover delay increases m by a given increment, Δm ; the greater the changeover delay the greater the value of Δm . Thus, with changeover delays:

$$\frac{I_1}{I_2} = \frac{m_1 + \Delta m}{m_2 + \Delta m} \cdot \frac{C_1}{C_2}. \quad (7)$$

If concurrent schedules have different values of m , exact matching will not be obtained. But adding a changeover delay will tend to promote matching, since greater m 's will cause the fraction, $(m_1 + \Delta m)/(m_2 + \Delta m)$ in Equation 7 to approach unity. In the computer sim-

ulation (Rachlin *et al.*, 1976) and in experiments with pigeons (Herrnstein, 1961), changeover delays tended to increase the tendency for matching to occur as Equation 7 predicts. The effect of the changeover delay, however, should be just the opposite with respect to the relation between responding and the *programmed* rates of reinforcement. Equation 5 predicts that even where $m_1 = m_2$, matching will not be obtained between I_1/I_2 and a_1/a_2 . As long as $m > 0$, the ratio of the responses will always differ from the ratio of programmed reinforcers. "Overmatching", according to Baum (1974), refers to a condition in Equation 5 where $m_1 = m_2 = m$ and $1-m$ is a fraction less than unity. The greater m is, the more overmatching there should be of relative responding to relative programmed reinforcements. Increasing the changeover delay, hence increasing m , should cause more overmatching with respect to programmed rates of reinforcement. This, again, was found in both the computer simulation (Rachlin *et al.*, 1976) and in experiments with rats (Shull and Pliskoff, 1967).

With concurrent ratio schedules, $m_1 = m_2 = 1$ and matching must occur according to Equation 6. Herrnstein (1970) showed that matching with unequal ratio schedules will be obtained only when either $I_1 = 0$ or $I_2 = 0$. In Equation 6, substituting aI for C :

$$\frac{I_1}{I_2} = \frac{a_1 I_1}{a_2 I_2}.$$

If a_1 is set unequal to a_2 the above equation will hold only if I_1 or I_2 equals zero. That is, responding will be all-or-none to one alternative. Such exclusive distribution was found in the computer simulation (Rachlin, *et al.*, 1976) and in experiments with pigeons (Herrnstein, 1961).

With concurrent variable-interval, variable-ratio (VI VR) schedules, the exponent, m , of the feedback function corresponding to the ratio schedule is 1.0, so Equation 6 becomes:

$$\frac{I_1}{I_2} = \frac{1}{m_2} \cdot \frac{C_1}{C_2},$$

where the subscript, 2, stands for the interval schedule. This equation is one for "biased" matching. Since m_2 is less than 1.0, I_1/I_2 should be greater than C_1/C_2 . In other words, subjects should spend proportionally more time on the ratio schedule than on the inter-

val schedule. If $m_2 \approx 0.2$, the bias would be 5:1 in favor of the ratio schedule. Staddon and Motherall (*in press*) predict a similar bias with a maximization analysis based on discrete rather than continuous responding. Contrary to these predictions, findings with concurrent VI VR schedules (experiments of Herrnstein and Loveland reported in deVilliers, 1977) are very close to unbiased matching (although variable). This discrepancy could be due to the presence of the changeover delay (COD). A COD would increase m_2 . If ratio-schedule responses accumulated when made during the COD, the COD would have no effect on m_1 , which would still equal unity. Thus, the COD should be a very important factor in VI VR choice experiments. With no COD, a strong bias should be found for the ratio schedule. With a COD, the bias should be reduced. To the author's knowledge, the COD has not been varied in VI VR experiments, so the prediction has yet to be tested.

Looking again at Equation 6, we see that matching is independent of the programmed rates of reinforcement, independent of the exponent, x , by which value of reinforcement increases with its duration, independent of the total duration of responding and even, in a sense, independent of m , the exponent of the feedback equation. Baum and Rachlin (1969) showed that a multiplicative term, such as m_1/m_2 in Equation 6, is properly interpreted as bias consistent with matching, rather than a deviation from matching. Given this universal independence, it is not surprising that matching is such a pervasive finding.² It is surprising rather that it is occasionally not found. Baum (1974) speculated that matching fails to occur when the source of each reinforcement (*i.e.*, which of the responses produced it) is obscured. Such an explanation would not fit with our assumption (implicit in Equation 4) that reinforcements from any source are equivalent. Deviations from biased matching, to be in accordance with our assumptions, would have to result from other factors. One such factor could be failure to measure accurately the reinforcement actually obtained. DeVilliers (1977) spec-

²Because matching is independent of the exponent, x , matching would follow from a function that equates value with reinforcement ($x = 1$, $y = 0$, $z = 0$). That is, matching follows from simple maximization of reinforcement.

ulated that the amount of reinforcement *actually* obtained (C) may be less than the amount measured (say due to loss of food from the hopper, delay in eating proffered food, failure to swallow the food eaten, failure to taste or chew all food swallowed, *etc.*).

Each delivered reinforcement is consumed with less than 100% efficiency. Some reinforcers will be actually consumed and some "wasted". Whether one finds matching, overmatching, or undermatching will depend on the relation of the efficiency to the amount delivered. If the amount wasted is proportionally greater for smaller amounts delivered, as it would be if the latency of pigeon's approach to a food-delivery hopper were a significant fraction of the delivery duration, overmatching will be obtained. If, on the other hand, the amount wasted is proportionally greater for larger amounts delivered, as it would be if a rat could not chew larger pellets or if a temporary satiation effect occurred with large local amounts or rates of reinforcement, undermatching will be obtained.

These relationships can best be illustrated by example. First, let us assume $m_1 = m_2$ in Equation 6. Suppose the amount delivered of C_1 is four units and the amount delivered of C_2 is six units. The relative rate of delivered reinforcement as measured is $4/10 = 0.4$. Now suppose the waste is one-fourth of the amount delivered, so the actual amounts obtained are three units and 4.5 units. The relative rate of obtained reinforcement (and hence of responding) is $3/7.5 = 0.4$, which matches the measured relative rate. Now suppose instead that the waste is a constant of one unit (proportionally greater for smaller values of C), so the actual amounts obtained are three of the four units delivered and five of the six units. The relative rate of obtained reinforcement (and, hence, of responding) is $3/8 = 0.375$, which is further from indifference than 0.4, the measured relative rate of reinforcement. This constitutes overmatching. With proportionally greater wastage for larger amounts, the one-fourth reduction of four units to three units might be accompanied by a one-third reduction of six units to four units. The relative rate of obtained reinforcement (hence, responding) would be $3/7 = 0.43$, which is closer to indifference than 0.4, the measured relative rate of reinforcement. This constitutes undermatching.

Another possible cause of undermatching, the deviation from matching most commonly (although not always) found (Baum, 1974) might be long-term order effects. Cross (1973) showed, with cross-modality matching in psychophysics, that regression effects are entirely due to the order in which the stimuli are presented, each response being a function of the present stimulus and the previous response. The effect occurs even with "random" orders of stimuli. A parallel phenomenon could occur in matching experiments; responding to a current condition could be a function of that condition plus the previous condition. Such "hysteresis" effects have been known to occur in choice experiments with pigeons (Rachlin, 1967) and may occasionally be responsible for undermatching. Thus, the degree of matching in an experiment may be a measure, not of the organism's tendency to match (which is an inevitable consequence of our assumptions) but of our ability to measure reinforcement actually obtained. I argued previously (Rachlin, 1971) that matching is best regarded as an assumption, rather than an empirical finding. Here we see that the assumption of matching is not inconsistent with the assumptions of the present argument.

Variable-Interval and Variable-Time Schedules

With variable-interval schedules, we assume that response rates are moderate and that I, the amount of responding, has little effect on the value of the session as a whole. Thus, y , in Equation 2 is zero and:

$$\begin{aligned} V_{(I,C,N)} &= V_C + V_I + V_N \\ &= C^x + I^0 + N^x \\ &= [aI^m]^x + 1 + [1 - I]^x \\ \frac{dV}{dI} &= mxa^x I^{mx-1} - z(1 - I)^{x-1}. \end{aligned}$$

Setting the derivative equal to zero:

$$mxa^x I^{mx-1} = z(1 - I)^{x-1}. \quad (9)$$

Substituting C for a from Equation 1 and rearranging terms:

$$C = \left[\frac{z}{mx} \cdot I(1 - I)^{x-1} \right]^{1/x}. \quad (10)$$

If we set $m = 0.1$ for typical variable-interval schedules, Equation 10 becomes an expression relating C, the observed reinforcement to I, the observed fraction of time responding on

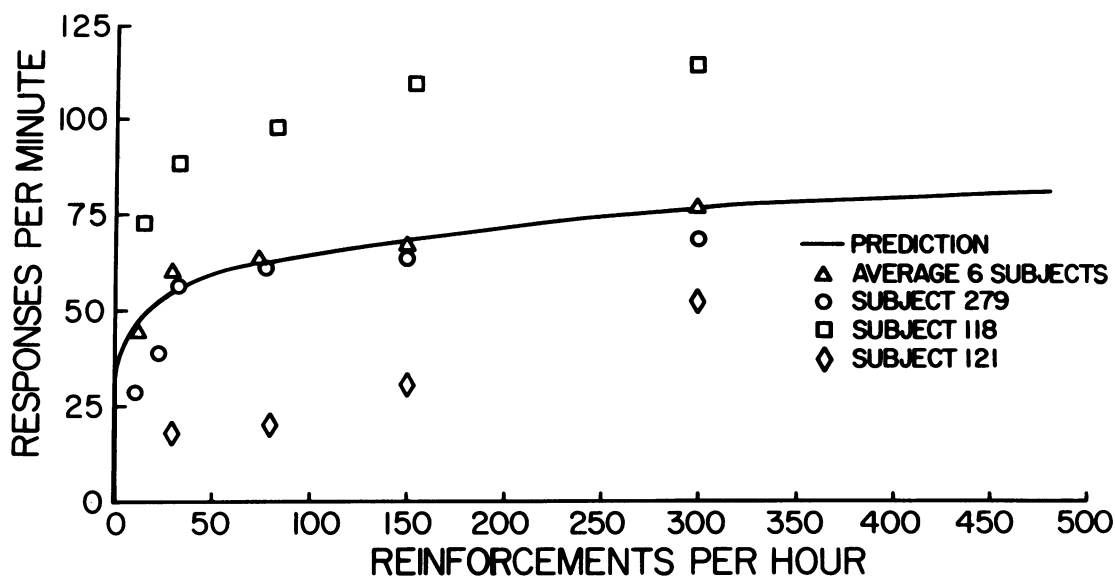


Fig. 2. Data from Catania and Reynolds (1968). Rate of responding as a function of rate of reinforcement in a variable-interval schedule. Solid line is predicted function.

a variable-interval schedule in terms of two constants, x and z . We can now take I and C from actual experiments with variable-interval schedules and try to determine the constants x and z . Catania and Reynolds (1968) exposed pigeons to various variable-interval schedules and measured rate of responding. Reinforcement duration was 4 sec when it was presented, so:

$$\left. \begin{aligned} C &= \frac{4 \text{ (seconds per reinforcement)}}{\text{observed seconds between reinforcements}} \\ I &= \text{observed responses per second} \times 0.25 \text{ sec per response} \end{aligned} \right\} \quad (11)$$

Appendix II of this paper describes how Equation 10 was fitted to the Catania-Reynolds data. Figure 2 shows predicted and actual functions for $m = 0.1$, $x = 0.2$, $z = 0.034$. In assuming that responding *per se* does not alter value we have assumed $y = 0$. Combining x , y , and z in Equation 10:

$$C = 14.2 \frac{I^3}{(1-I)^{4.83}}.$$

The solid line shows the relationship between reinforcements per hour and responses per minute predicted by the above equation. The triangles are averages of six pigeons tested by Catania and Reynolds. The diamonds, circles, and squares are data of the slowest-responding pigeon, an average pigeon, and the fastest-responding pigeon. The solid line fits the data

about as well as a single line could fit such variable data.

It is not surprising that an equation with several constants can be made to fit the results of a given experiment. A better test of the model is whether the same constants can be used to explain results of other experiments. Before turning to this question, however, let us consider another way to represent the data, which better illustrates the dynamics of the model. We have assumed that $I + N = 1$ represents a unit session. Thus, a large C would result in a longer actual session. To normalize C , I , and N with respect to actual session duration:

$$\left. \begin{aligned} \bar{C} &= \frac{C}{I+N+C} = \frac{C}{1+C} \\ \bar{I} &= \frac{I}{I+N+C} = \frac{I}{1+C} \\ \bar{N} &= \frac{N}{I+N+C} = \frac{N}{1+C} \end{aligned} \right\} \quad (12)$$

The response-variables \bar{C} , \bar{I} , and \bar{N} add to 1.0 and may be represented as corners of the equilateral triangle in Figure 3, (Rachlin and Burkhard, 1978; Staddon, *in press*). The point, $\bar{C} = 1$, $\bar{I} = 0$, $\bar{N} = 0$ is at the \bar{C} -vertex of the triangle; the point, $\bar{C} = 0$, $\bar{I} = 1$, $\bar{N} = 0$ is at the \bar{I} -vertex; the point, $\bar{C} = 0$, $\bar{I} = 0$, $\bar{N} = 1$ is at the \bar{N} -vertex. A point on an edge of the triangle represents zero time allocated to the response on the opposite vertex, and an alloca-

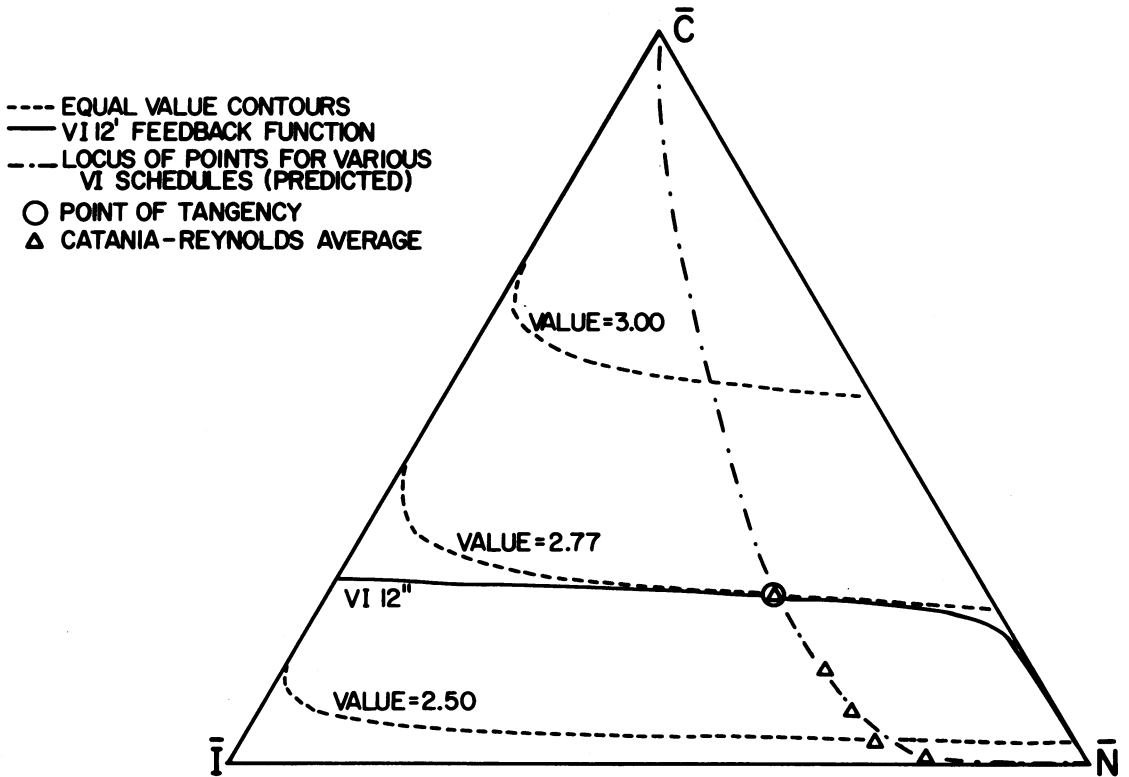


Fig. 3. Representation of $\bar{C} + \bar{I} + \bar{N} = 1.0$ on triangular coordinates. Solid line is feedback function for VI 12-sec schedule. Dotted lines are equal-value contours: $V = C^{0.2} + I^{0.098}$. Circle is point of tangency of solid with dotted line. Broken line is locus of such points of tangency for different variable-interval schedules. Small triangles are average data from Catania and Reynolds (1968).

tion of time to the other two responses in proportion to the closeness of the point to their respective vertices. Any point in the triangle represents a given allocation of behavior among the three responses determined by its distance from the three vertices. If a line were drawn in the triangle from one vertex to the opposite side, it would contain all points of a fixed proportion of the two responses corresponding to the two opposite vertices. Thus, a line from the \bar{C} -vertex to the \bar{I} - \bar{N} edge bisecting the \bar{I} - \bar{N} edge would represent all points at which $\bar{I} = \bar{N}$. The point at the \bar{C} -vertex would be $\bar{I} = \bar{N} = 0, \bar{C} = 1$. The point where the line would intersect the \bar{I} - \bar{N} edge would be $\bar{I} = \bar{N} = 0.5, \bar{C} = 0$. The point in the middle of the triangle is $\bar{I} = \bar{N} = \bar{C} = 0.33$. All possible allocations of behavior to \bar{I}, \bar{N} , and \bar{C} are represented in the triangle.³

In the triangle of Figure 3, the dotted lines are contours of equal value according to Equation 13:

$$\begin{aligned} V &= C^x + I^y + N^z \\ V &= C^{0.2} + I^0 + N^{0.098} \\ V &= C^{0.2} + 1 + N^{0.098}, \end{aligned} \tag{13}$$

where C and N are converted to \bar{C} and \bar{N} in the figure. The dotted lines shown are samples from an infinite set of lines of similar shape ranging from low values low in the triangle to high values high in the triangle. The highest point is at the \bar{C} -vertex. The lowest point is at the \bar{I} -vertex. The solid lines represent the feedback equation (Equation 1) for a variable-interval schedule with 12 sec programmed between reinforcements (VI 12-sec). The feedback function restricts the possible allocations (each allocation represented by a point in the triangle) to those on (or below) the solid line.

³Triangular coordinate graph paper is used in chemistry to represent various compounds in solution. It is available from most companies that make engineering graph paper (e.g., #12-082, National Blank Book Company, Holyoke, Massachusetts 01040).

Feedback functions of higher VI schedules (e.g., VI 1-min) would be lower in the triangle. Those of lower VI schedules (e.g., VI 8-sec) would be higher. We assume that the subject will allocate behavior (among \bar{C} , \bar{I} , and \bar{N}), so as to maximize value. Thus, the subject must respond at the point on the feedback function that touches the contour of highest value. This point will be where the feedback function is tangent to the highest possible value contour. In Figure 3, the point of tangency is shown as a circle and the locus of such points with various other variable-interval schedules is shown as a broken line. The data of Catania and Reynolds (1968) converted to appropriate units by Equations 11 and 12 are plotted as small triangles.

The fact that the Catania-Reynolds data was fit by the model indicates that a single set of value contours can apply with various variable-interval schedules. Variable-time (VT) schedules are described by feedback functions that, in Figure 3, would be horizontal lines (constant \bar{C}). Because the value contours are all sloped slightly upward to the left, a horizontal feedback function would touch the highest value contour on the \bar{C} - \bar{N} edge (where $\bar{I} = 0$). Thus, \bar{I} should equal zero with variable-time schedules, corresponding to the finding that free reinforcement at variable intervals does not generate responding. However, the shallow slope of the variable-interval feedback function and a horizontal function might well prove difficult to discriminate once responding had reached the value it would attain with a variable-interval schedule. Thus, the model can account for the "superstitious" persistence of responding found when schedules are changed from VI to VT.

Note that in the broken line of Figure 3, \bar{I} first increases and then decreases as \bar{C} increases from zero (the \bar{I} - \bar{N} edge) to 1.0 (the \bar{C} -vertex). However, in Figure 2, rate of response increases continuously with rate of reinforcement. This is because of the transformation from I in Figure 2 to \bar{I} in Figure 3. At relatively high values of C , \bar{I} will be much smaller than I (as Equation 12 demands).

Variable-Ratio Schedules

Feedback functions for variable-ratio (VR) schedules are lines of constant proportionality between C and I and (as can be seen from Equation 12) of an equal proportionality be-

tween \bar{C} and \bar{I} . Such a line on the triangle of Figure 3 is a straight line from the \bar{N} -vertex to edge \bar{C} - \bar{I} . High variable-ratio schedules would have flatter slopes. A line along the \bar{N} - \bar{I} edge corresponds to $VR \infty$ (i.e., extinction); a line along the \bar{N} - \bar{C} edge corresponds to $VR 0$ (C , continuously available; I , unavailable). It is obvious from Figure 3 that points of tangency of straight lines (from \bar{N} -vertex to edge \bar{C} - \bar{I}) and the value contours would be generally to the left of those of VI feedback functions. Figure 4 shows two examples of VR feedback functions (still assuming each response takes 0.25 sec and each reinforcement lasts 4 sec), $VR 10$ and $VR 100$ and the value contours (using the same constants as in Figure 3) to which they are tangent. The actual points of tangency were determined by the method described in Appendix II. The broken line is the locus of points of tangency from $VR 0$ to $VR \infty$. Figure 5 shows this function transformed into a curve of responses per minute *versus* the VR schedule. The function involves higher rates of response than that of Figure 2, as would be expected with ratio schedules.

There is a paucity of actual data to compare with these predictions, but those data that are available with pigeon subjects, fixed-ratio functions obtained by Felton and Lyon (1966) and variable-ratio functions in multiple schedules obtained by Zuriff (1970), are lower in rate of response than the curve of Figure 5 (although still higher than the rates of response obtained with interval schedules) and fall off suddenly in rate at high ratios ("ratio-strain"). Although the function of Figure 5 decreases with increased ratios, it does not decrease suddenly as ratios increase, at least within the range of ratios normally programmed.

At this point, we have to abandon the last of the simplifying assumptions made at the beginning: that responding is of moderate rate and in itself has no effect on the value of the session. With interval schedules, the assumption probably is valid, at least with pigeons pecking keys. However, with ratio schedules, and the high response rates they generate, even small effects of each response on value would cumulate significantly. It seems likely that with low rates of response, responding is insignificant in value, but with high rates of response, responding *per se* has a negative in-

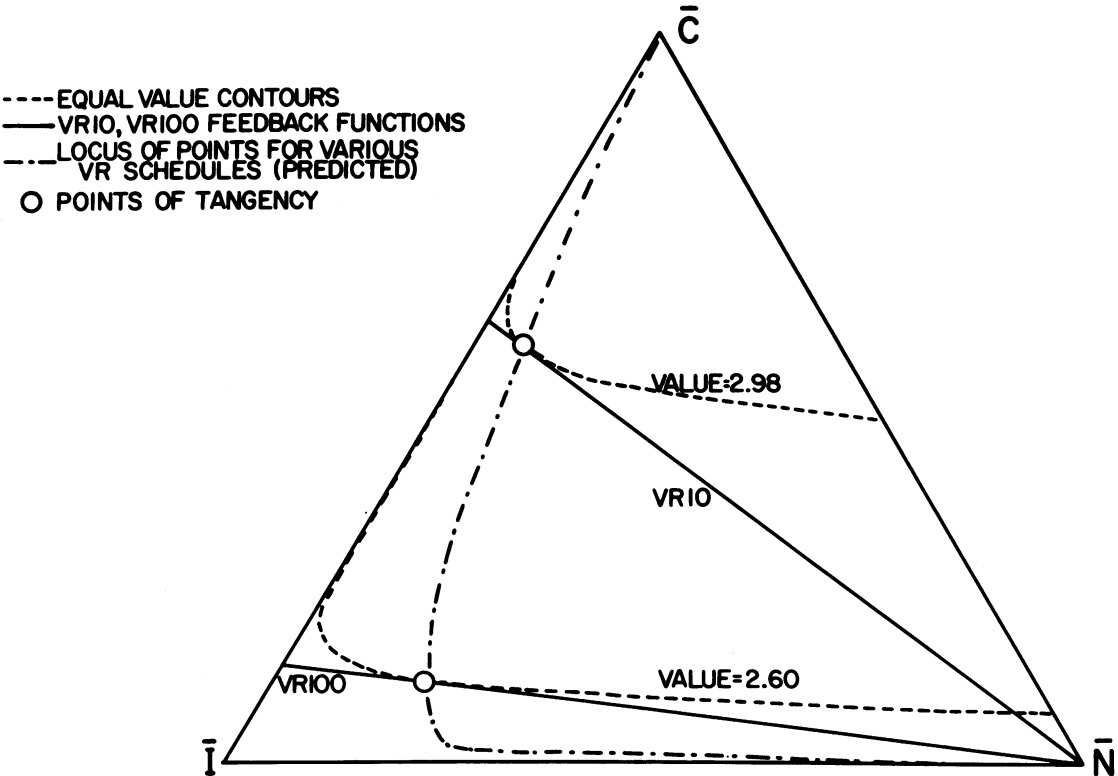


Fig. 4. Representation of $\bar{C} + \bar{I} + \bar{N} = 1.0$ as in Figure 3. Solid lines are feedback functions for variable-ratio schedules. Again, $V = C^{0.9} + N^{0.099}$.

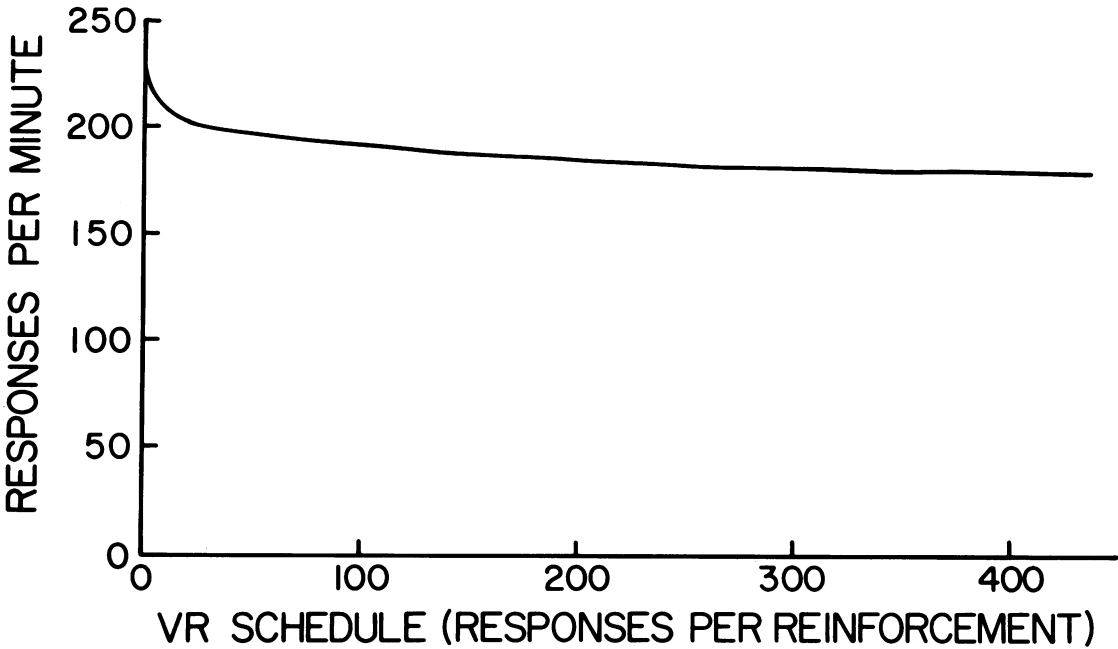


Fig. 5. Rate of responding as a function of the variable-ratio schedule imposed. This function is the broken line of Figure 4 converted from temporal units to rate units.

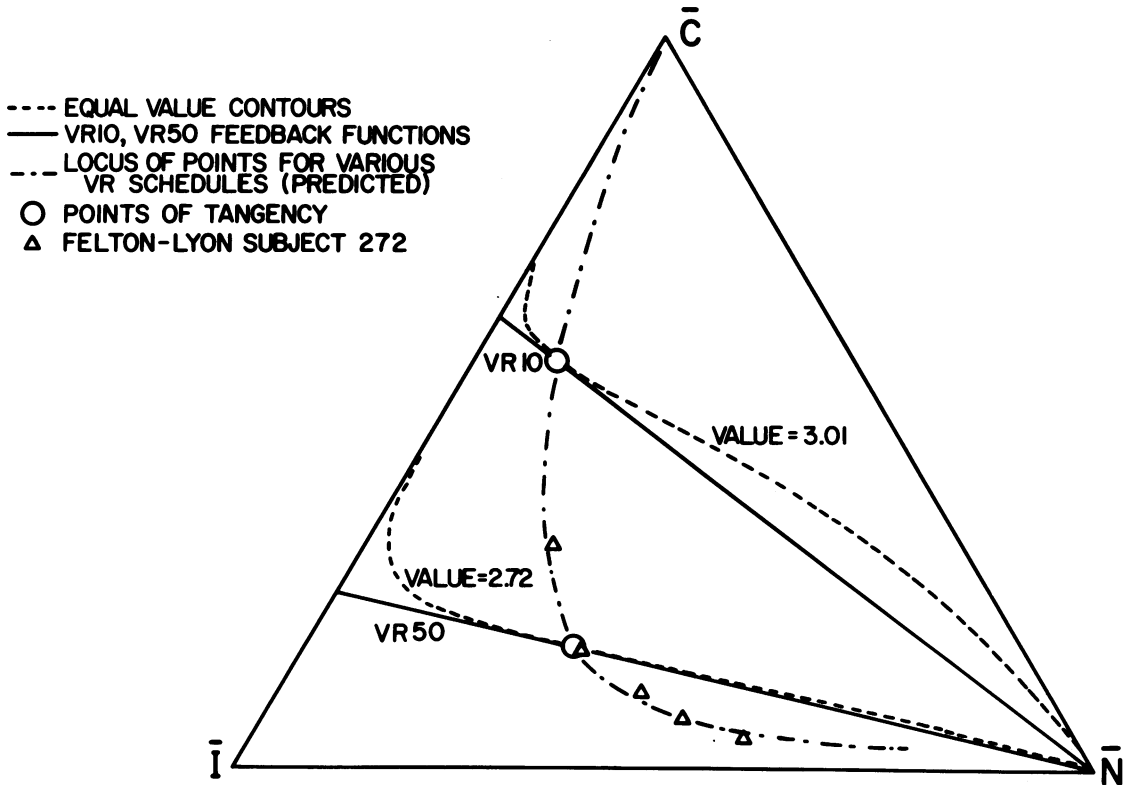


Fig. 6. Representation of $\bar{C} + \bar{I} + \bar{N} = 1.0$ as in Figure 4. Value contours are different in this figure because here $V = C^{0.2} + I^{-0.08} + N^{0.038}$. Small triangles are data from Felton and Lyon (1966).

fluence on value. That is, high rates of response are assumed to be slightly punishing. This punishing effect can be incorporated into the model by abandoning the assumption that, in Equation 2, $y = 0$ and making y negative. Figure 6 illustrates the effects of assigning $y = -0.08$. The particular exponent, -0.08 was chosen, by the method described in Appendix II, to fit the data of a typical one of Felton and Lyon's subjects (#272). The negative sign of y bends the value contours inward. The equation for value is now:

$$V = C^{0.2} + I^{-0.08} + N^{0.038} \quad (14)$$

Two VR-schedule feedback functions, VR 10 and VR 50, are tangent to the two value contours shown at points where circles are drawn in Figure 6. The locus of points of tangency is shown as a broken line in Figure 6. Note that as the VR schedule increases, this locus falls away from the \bar{I} -vertex much more rapidly than the corresponding locus of Figure 4 (where $y = 0$). The small triangles are data of Felton and Lyon's Subject #272. Note that

because of the inward bending of the value contours each ratio feedback function touches a given value contour at two points: the point on the broken line and the point at the \bar{N} -vertex (where $\bar{I} = \bar{C} = 0$). The "true" value contours probably look like those of Figure 4 when \bar{I} is small and like those of Figure 6 when \bar{I} is large. Note also that the broken line of Figure 6 becomes nearly horizontal above the \bar{I} - \bar{N} edge of the triangle. High ratio schedules (above FR 130) do not intersect this line, and thus touch the value contours of Figure 6 *only* at the \bar{N} -vertex. Thus, responding with such high-valued ratio schedules should be unstable. At $I = 0$, the value contours of Figure 4 should apply and responding at high rates should occur. But as soon as responding actually does increase sufficiently, the value contours of Figure 6 should apply and responding should decrease to zero. The model thus accounts for the erratic behavior at very high ratios (ratio strain). Figure 7 converts and replots the broken line of Figure 6 on the same axes as Figure 5. The triangle

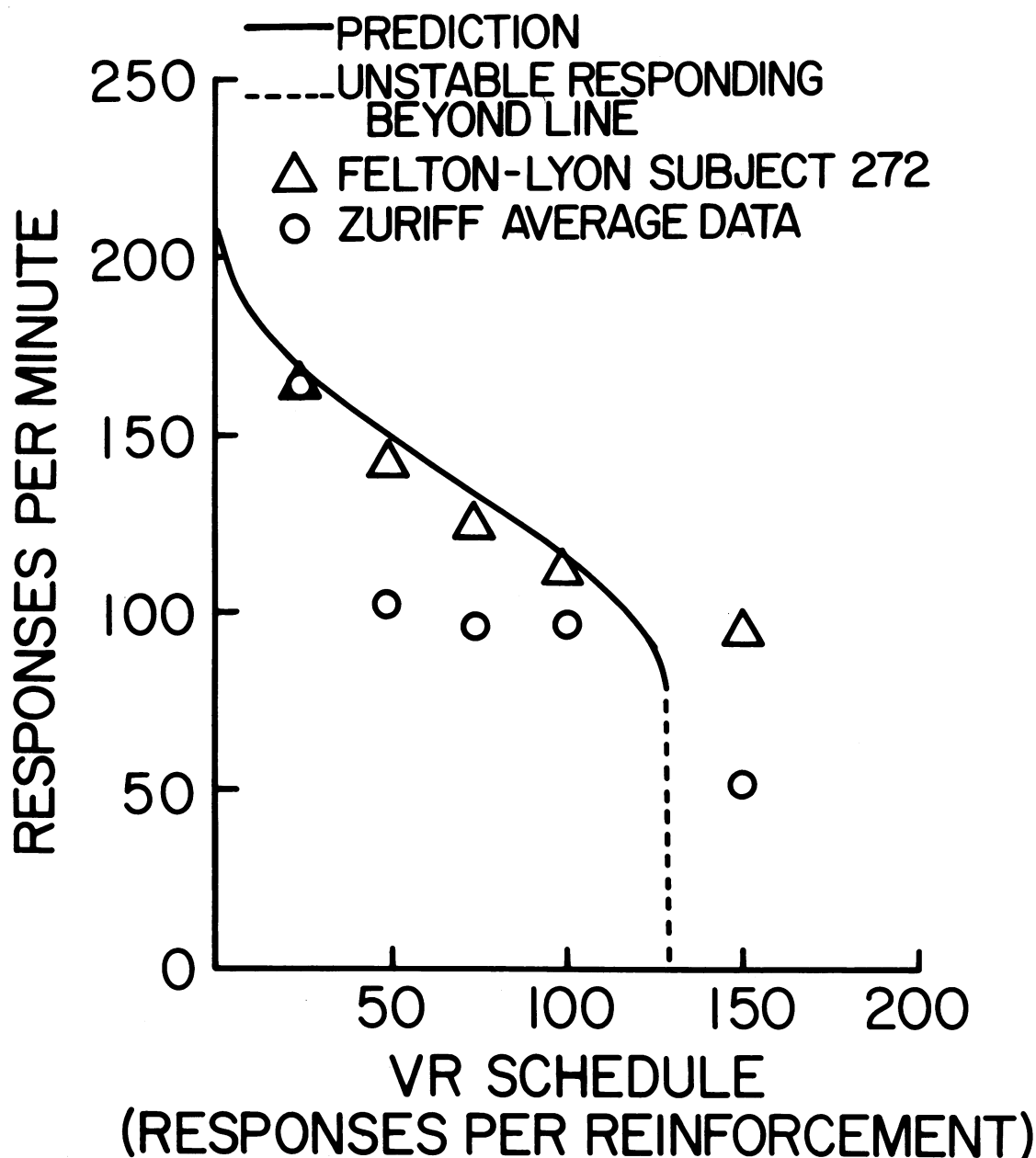


Fig. 7. Rate of responding as a function of the variable-ratio schedule imposed. Here, the broken line of Figure 6 is converted from temporal units to rate units. The triangles are data of Felton and Lyon (1966). The circles are data of Zuriff (1970). The dotted portion of the curve represents unstable responding (ratio strain).

points again are the data of Felton and Lyons' subject. The circled points are the data of an average subject of Zuriff (1970).

This method of fitting ratio and interval data is probably not the most efficient possible. An alternative would have been to consider both interval and ratio data together, and arrive at a common set of exponents x , y , and z

in the equation for value. Still another alternative would have been to introduce coefficients for the terms in the equation. The negative exponent creates some awkwardness at boundary conditions and furthermore violates some reasonable assumptions about behavior (diminishing marginal substitutability of responses) that are inconsistent with the inward

bend of the contours of Figure 6 (Rachlin and Burkhard, 1978). Nevertheless, the ease with which the value equation can account for the schedule data encourages the general assumption that subjects exposed to schedules of reinforcement adjust their behavior so as to maximize its value within the restrictions that the schedule imposes.

CONCLUSION

Behavior under some commonly studied schedules of reinforcement is not inconsistent with the assumption that value grows as a power function of response duration and that subjects allocate behavior among responses so as to maximize value. No proof of these assumptions is offered or suggested. Their truth or falsity will rest on the stability of the constants in the equations, the development of methods to evaluate the various constants (all of which may be independently determined), and the extension of the model to other areas. A more general form of the model, with determinations of value other than those of Equations 2 and 3, is described in another article (Rachlin and Burkhard, 1978).

APPENDIX I

Derivation of the feedback function. Suppose the pattern of responses is alternation of bursts and pauses. We can consider a cycle of duration T minutes divided into a period of responding, with expected value i minutes, and a period of nonresponding, with expected value n minutes, so that $i + n = T$. A session, S , consists of a number, B , of cycles so that $S = B \cdot T$. The variable-interval schedule, when it is running, delivers reinforcements at a rate of r reinforcements per minute.

Our first object is to determine the expected number of reinforcements during a cycle, T . The variable-interval schedule is running continuously during interval i and delivering reinforcements, so that the expected number of reinforcements during i is ri . During n , the variable-interval schedule will run only until a reinforcement sets up. If a reinforcement is set up, it will not be delivered until the beginning of the next i -period. Nevertheless, we assign the reinforcement to the n -period during which it is set up. Since, at most, one reinforcement can set up during an n -period, the expected number of reinforcements must be less

than 1. It is equal to the probability of a reinforcement setting up during the n -period, which is 1 minus the probability of no reinforcements setting up. If, as is usually the case, the variable intervals are exponentially distributed, the probability of no reinforcement setting up is e^{-rn} . Thus, the probability of a reinforcement setting up during an n -period is $1 - e^{-rn}$ and the expected number of reinforcements during a cycle, T , is $ri + (1 - e^{-rn})$.

The expected number of reinforcements during a session is B , the number of cycles, multiplied by the number of reinforcements per cycle (assuming fixed cycle duration), or $B \cdot (ri + 1 - e^{-rn})$.

The variable, C (proportion of session spent consuming reinforcement), is this expected number times the duration of each reinforcement (H , for "hopper-time-up") divided by the session duration:

$$C = \frac{BH}{S} (ri + 1 - e^{-rn}).$$

The variable, I (proportion of session responding) is, similarly, the period, i , of the cycle spent responding times the number of cycles divided by the session duration:

$$I = \frac{Bi}{S}.$$

Now, n , the period spent not responding, equals $T - i$ and $T = S/B$ so

$$n = \frac{S}{B} - \frac{Si}{B} = \frac{S}{B} (1 - I).$$

Substituting, I , S , and B for T , i , and n :

$$C = \frac{HB}{S} \left[r \frac{Si}{B} + 1 - e^{-\frac{rS}{B} (1-I)} \right]$$

$$C = Hri + \frac{HB}{S} \left[1 - e^{-\frac{rS}{B} (1-I)} \right]. \quad (15)$$

The constant, a , the programmed proportion of reinforcement is equal to rH . Substituting in Equation 1:

$$C = aI + \frac{HB}{S} \left[1 - e^{-\frac{aS}{BH} (1-I)} \right]. \quad (16)$$

This is the feedback function for variable-interval schedules. The function is somewhat awkward to use. Figure 8 plots C versus I on log-log coordinates with various values of B (the number of cycles) assuming a 60-min session ($S = 60$) and a 4-sec reinforcement de-

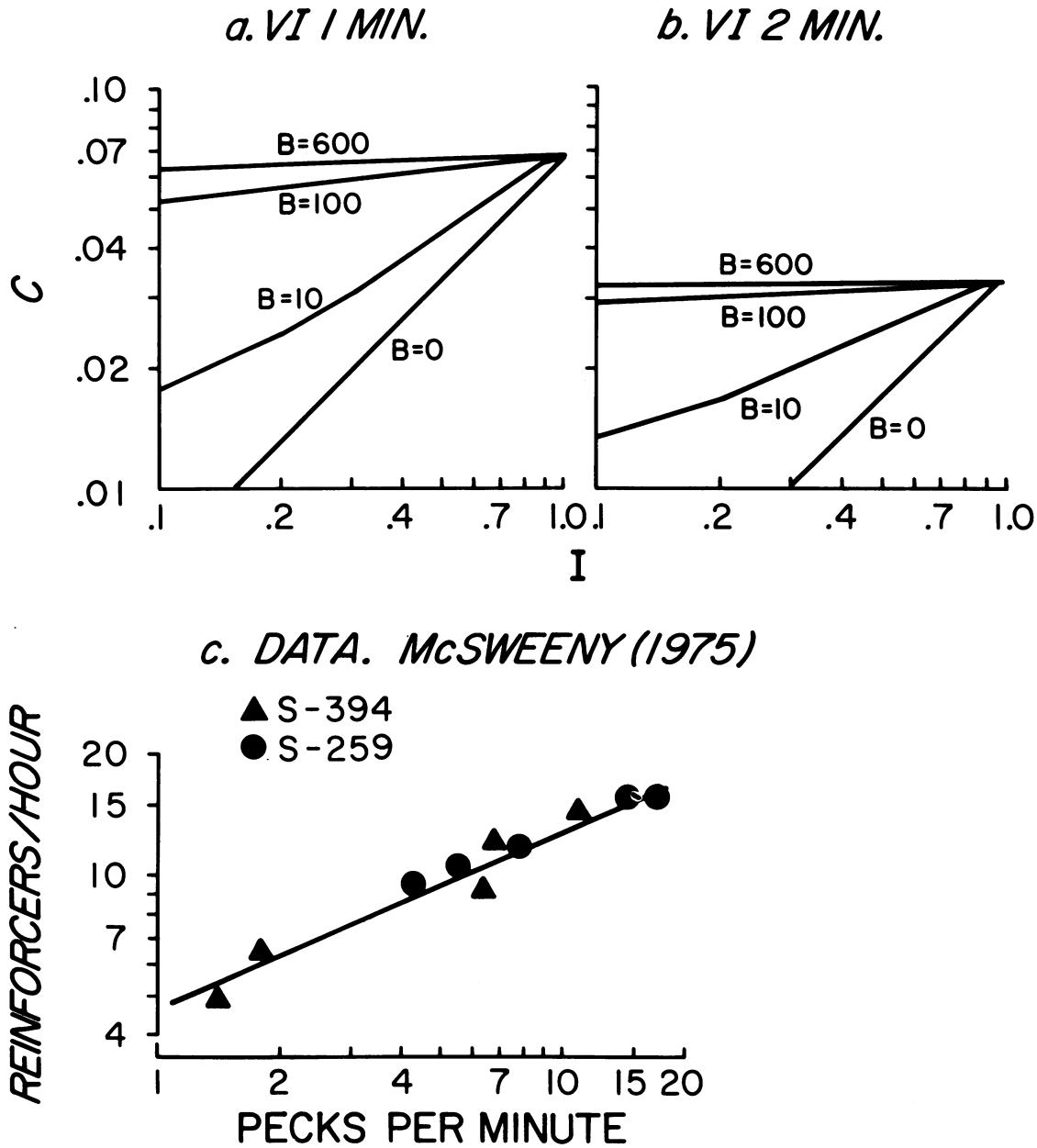


Fig. 8. a. Equation 16, the feedback equation derived in Appendix I for a VI 1-min schedule. Reinforcement (C) is a function of responding (I). The parameter (B) is the number of burst-pause cycles occurring in a 60-min session. b. Same as a, but for a VI 2-min schedule. c. An empirical function obtained with a VI 4-min schedule. The ordinate and abscissa may be converted to C and I by multiplication by appropriate constants. This would not change the shape of the function.

livery ($H = r/60$) with two variable-interval schedules, VI 1-min ($r = 1$ reinforcement per minute; $a = 0.067$) and VI 2-min ($r = 0.5$ reinforcements per minute; $a = 0.033$). When B is low, C is low because reinforcements are lost when long periods of pausing (high ab-

solute values of n) intervene between periods of responding. Although the functions are slightly S-shaped, they approximate straight lines and are well approximated by the simple power function of Equation 1 (straight lines on log log coordinates are power functions).

The slope of the function is determined by B, the number of cycles per session. This constant corresponds to the constant, m, in Equation 1. When a COD is added, switches are penalized and B is in fact reduced. This has the effect of increasing m, the slope of the line in Figure 8. As indicated in the text, equal increases of m for two schedules tend to reduce bias and promote matching. Thus, higher CODs promote matching.

For convenience, we assumed in the text that m was constant for various variable-interval schedules. But equal values of B give different slopes for different schedules as indicated by Figure 8. In order for the slope (m) to remain constant, B would have to change; subjects would have to alternate between responding and not-responding less frequently with lower rates of reinforcement. Informal observation of cumulative records indicates that pigeons do pause longer with longer VI schedules. If, however, m were different for different VI schedules, biased matching would be obtained and the function of Figure 2 would be altered. It is thus an important empirical question whether m stays approximately constant across VI schedules. This question has not yet been addressed in experiments.

Figure 8 also shows an actual feedback function from data obtained by McSweeney (1975) with concurrent variable-interval responding of pigeons at various body weights. The two sets of points are for the only two (of five) pigeons for which a complete set of body weights was obtained. The schedule shown (VI 4-min) was the lowest rate of reinforcement used. With higher rates of reinforcement, response rates tended to bunch at high values. The spread of points shown in Figure 8 is approximated by a straight line on log log coordinates. But note that the points show considerable fluctuation, indicating that the feedback function is sensitive to the patterning of responding.

At $B = 0$, $C = aI$, which is the feedback function for variable-ratio schedules. The feedback function has no meaning where $I = 0$ and $B = 0$. This corresponds to alternation of non-responding (n) with an i of zero. The reinforcement set up in n would never be collected. Thus, when I equals zero, B must also equal zero. At this point (and near this point) the VI and VR schedules are equivalent.

APPENDIX II

Fitting the Catania-Reynolds (1968) data. Equations 9 and 10 are two versions of the same equation; one, in terms of programmed rate of reinforcement; the other, in terms of actual rate. We fit Equation 9, with the programmed rates of reinforcement, since these could be calculated and did not need to be estimated from Figure 1 in the Catania-Reynolds paper, as did rates of response. We assume $m = 0.1$. From Equation 11 (a instead of C) with a VI 1-min schedule, $a = 4/60 = 0.067$. From Equation 11 and Figure 1 (of Catania-Reynolds, 1968), with a VI 1-min schedule for an average pigeon (rate of response is 60 per minute = 1 per second) assuming each response takes 0.25 sec, $I = 1 \times 0.25 = 0.25$. Substituting $m = 0.1$, $a = 0.067$, $I = 0.25$ in Equation 9:

$$0.1x(0.067)^x(0.25)^{-1x-1} = z(0.75)^{x-1}.$$

Doing the equivalent for VI 12-sec:

$$0.1x(0.33)^x(0.32)^{-1x-1} = z(0.68)^{x-1}.$$

Solving the two simultaneous equations for x and Z:

$$x = 0.2, z = 0.034.$$

With ratio schedules. Now we no longer assume $y = 0$. Thus, from Equations 2 and 3:

$$V = C^x + I^x + N^x.$$

From Equation 1 and the relation, $N = 1 - I$,

$$V = (aI^m)^x + I^x + (1 - I)^x$$

$$\frac{dV}{dI} = mxa^x I^{mx-1} + yI^{x-1} - z(1 - I)^{x-1}.$$

In ratio schedules, $m = 1$. From the interval schedule analysis, $x = 0.2$, $z = 0.038$. From Zuriff's (1970) data, with VR 100, rate of response was about 100 responses per minute. Thus, $a = C/I = 4/(100 \times 0.25) = 0.16$ and $I = (100 \times 0.25)/60 = 0.42$ (using conversion factors in Equation 11). Substituting these values into the derivative above and setting it equal to zero:

$$0 = 0.2(0.16)^{0.2}(0.42)^{-0.8} + y(0.42)^{x-1} - 0.038(0.58)^{-0.992}$$

Solving the above, $y = -0.08$.

REFERENCES

- Anger, D. The dependence of interresponse times upon the relative reinforcement of different inter-

- response times. *Journal of Experimental Psychology*, 1956, 52, 145-161.
- Baum, W. M. The correlation based law of effect. *Journal of the Experimental Analysis of Behavior*, 1973, 20, 137-153.
- Baum, W. M. On two types of deviation from the matching law: bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 1974, 22, 231-242.
- Baum, W. M. and Rachlin, H. C. Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, 1969, 12, 851-874.
- Catania, A. C. and Reynolds, G. S. A quantitative analysis of responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 327-383.
- Cross, D. V. Sequential dependencies and regression in psychophysical judgments. *Perception and Psychophysics*, 1973, 14, 547-552.
- deVilliers, P. Choice in concurrent schedules and a quantitative formulation of law of effect. In W. K. Honig and J. E. R. Staddon, (Eds), *Handbook of operant behavior*. Englewood Cliffs, New Jersey: Prentice-Hall, 1977. Pp. 233-287.
- Felton, M. and Lyon, D. O. The post-reinforcement pause. *Journal of the Experimental Analysis of Behavior*, 1966, 9, 131-135.
- Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1961, 4, 267-272.
- Herrnstein, R. J. On the law of effect. *Journal of the Experimental Analysis of Behavior*, 1970, 13, 243-266.
- McGill, W. J. Stochastic latency mechanism. In R. D. Luce, R. R. Bush, and E. Galanter (Eds), *Handbook of mathematical psychology*, Vol. 1. New York: Wiley, 1963. Pp. 309-361.
- McSweeney, F. K. Concurrent schedule responding as a function of body weight. *Animal Learning and Behavior*, 1975, 3, 264-270.
- Premack, D. Reinforcement theory. In D. Levine (Ed), *Nebraska symposium on motivation*: 1965. Lincoln: University of Nebraska Press, 1965. Pp. 123-188.
- Rachlin, H. The effect of shock intensity on concurrent and single-key responding in concurrent chain schedules. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 87-93.
- Rachlin, H. On the tautology of the matching law. *Journal of the Experimental Analysis of Behavior*, 1971, 15, 249-251.
- Rachlin, H. and Burkhard, B. The temporal triangle: Response substitution in instrumental conditioning. *Psychological Review*, 1978, 85, 22-48.
- Rachlin, H., Green, L., Kagel, J. H., and Battalio, R. C. Economic demand theory and psychological studies of choice. In G. Bower (Ed), *The psychology of learning and motivation*, Vol. 10. New York: Academic Press, 1976. Pp. 129-154.
- Shimp, C. P. Reinforcement of short interresponse times. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 425-434.
- Shull, R. L. and Pliskoff, J. J. Changeover delay and concurrent schedules: some effects on relative performance measures. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 517-529.
- Staddon, J. E. R. Operant behavior as adaptation to constraint. *Journal of Experimental Psychology: General* (in press).
- Staddon, J. E. R. and Motherall, S. On matching and maximizing in operant choice experiments. *Psychological Review* (in press).
- Zuriff, G. E. A comparison of variable-interval and variable-ratio schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1970, 13, 369-375.

Received 9 January 1978.

(Final acceptance 25 May 1978.)